

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of :
Antonio COLMENAREZ et al :
Serial No. n/a : Group Art Unit:
Filed: Herewith : Examiner:
For: AUTOMATIC SPEECH RECOGNITION SYSTEM AND METHOD

PRELIMINARY AMENDMENT

Assistant Commissioner for Patents
Washington, D. C. 20231

Sir:

Prior to examination, please amend the referenced application
as follows:

IN THE CLAIMS:

Please amend claims 1, 9, 13-18 and 25-27 as follows:

1. (amended) A speech recognition system comprising an
acoustic detector for detecting speech utterances of a speaker; a
visual detector for detecting at least one facial characteristic
associated with speech utterances of the speaker; a processing
arrangement connected to be responsive to the acoustic and visual
detectors for deriving a signal having first and second values
respectively indicative of the speaker making and not making speech
utterances such that the first value is derived in response to the

Serial No. sn

acoustic detector detecting a finite, nonzero acoustic response while the visual detector detects at least one facial characteristic associated with speech utterances of the speaker; and a speech recognizer for deriving an output indicative of the speech utterances as detected only by the acoustic detector, the speech recognizer being connected to be responsive to the acoustic detector in response to the signal having the first value.

9. (amended) The speech recognition system of claim 1 wherein the processing arrangement includes a delay arrangement for assuring that in response to completion of each speech utterance the acoustic detector is decoupled from the speech recognizer.

13. (amended) The speech recognition system of claim 12 wherein the face recognizer is arranged for enabling the signal to have the first value in response to the face of the speaker being at a predetermined orientation relative to the visual detector.

14. (amended) The speech recognition system of claim 13 wherein the face recognizer is arranged for: (1) detecting and distinguishing the faces of a plurality of speakers, and (2) enabling the signal to have the first value in response to the speaker having a recognized face.

15. (amended) The speech recognition system of claim 14 wherein the processing arrangement includes a speaker identity recognizer connected to be responsive to the acoustic detector, the speaker identity recognizer being arranged for: (1) detecting and

Serial No. sn

distinguishing speech patterns of a plurality of speakers, and (2) enabling the signal to have the first value in response to the speaker having a recognized speech pattern.

16. (amended) The speech recognition system of claim 15 wherein the processing arrangement is arranged for causing the signal to have the first value in response to the speaker having a recognized face matched with a recognized speech pattern of the same speaker.

17. (amended) The speech recognition system of claim 1 wherein the processing arrangement includes a face recognizer connected to be responsive to the visual detector and a speaker identity recognizer connected to be responsive to the acoustic detector, the face recognizer being arranged for detecting and distinguishing the faces of a plurality of speakers, the speaker identity recognizer being arranged for detecting and distinguishing speech patterns of a plurality of speakers, the processing arrangement being arranged for enabling the signal to have the first value in response to the speaker having a recognized face matched with a recognized speech pattern of the same speaker.

18. (amended) A method of recognizing speech utterances of a speaker with an automatic speech recognizer only responsive to acoustic speech utterances of the speaker comprising detecting acoustic energy having a spectrum associated with speech utterances, detecting at least one facial characteristic associated

Serial No. sn

with speech utterances of the speaker, and activating the automatic speech recognizer in response to the detected acoustic energy having a spectrum associated with speech utterances while the at least one facial characteristic associated with speech utterances of the speaker is occurring.

25. (amended) The method of claim 24 further comprising distinguishing the face of the speaker from a plurality of speakers, distinguishing the speech pattern of the speaker from a plurality of speakers, and activating the automatic speech recognizer in response to the speaker having a recognized face matched with a recognized speech pattern of the same speaker.

26. (amended) The method of claim 18 further comprising distinguishing the face of the speaker from a plurality of speakers, distinguishing the speech pattern of the speaker from a plurality of speakers, and activating the automatic speech recognizer in response to the speaker having a recognized face matched with a recognized speech pattern of the same speaker.

27. (amended) The method of claim 26 further including storing: (1) images of the faces of a plurality of speakers, and (2) the speech patterns of the same plurality of speakers during at least one training period; and performing the distinguishing step by comparing the stored images and speech patterns with images of the face of the speaker and the speech pattern of the speaker.

Serial No. sn

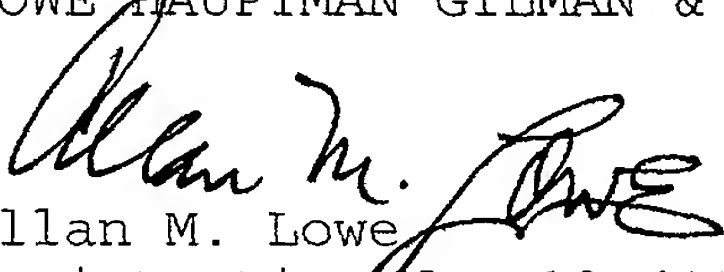
R E M A R K S

Upon further review of the claims of the present application, it was realized that some of the claims should be broadened somewhat and were possibly misdescriptive. The present Amendment cures these problems.

Entry of the Amendment is in order.

Respectfully submitted,

LOWE HAUPTMAN GILMAN & BERNER, LLP


Allan M. Lowe
Registration No. 19,641

1700 Diagonal Road, Suite 310
Alexandria, Virginia 22314
(703) 684-1111/FAX: (703) 518-5499
AML:ssw
Date: January 29, 2002

MARKED-UP VERSION SHOWING CHANGES

1. (amended) A speech recognition system comprising an acoustic detector for detecting speech utterances of a speaker; a visual detector for detecting at least one facial characteristic associated with speech utterances of the speaker; a processing arrangement connected to be responsive to the acoustic and visual detectors for deriving a signal having first and second values respectively indicative of the speaker making and not making speech utterances such that the first value is derived [only] in response to the acoustic detector detecting a finite, nonzero acoustic response while the visual detector detects at least one facial characteristic associated with speech utterances of the speaker; and a speech recognizer for deriving an output indicative of the speech utterances as detected only by the acoustic detector, the speech recognizer being connected to be responsive to the acoustic detector [only while the signal has] in response to the signal having the first value.

9. (amended) The speech recognition system of claim 1 wherein the processing arrangement includes a delay arrangement for assuring that [upon the] in response to completion of each speech utterance the acoustic detector is decoupled from the speech recognizer.

13. (amended) The speech recognition system of claim 12 wherein the face recognizer is arranged for enabling the signal to have the first value [only] in response to the face of the speaker being at a predetermined orientation relative to the visual detector.

14. (amended) The speech recognition system of claim 13 wherein the face recognizer is arranged for: (1) detecting and distinguishing the faces of a plurality of speakers, and (2) enabling the signal to have the first value [only] in response to

Serial No. sn

the speaker having a recognized face.

15. (amended) The speech recognition system of claim 14 wherein the processing arrangement includes a speaker identity recognizer connected to be responsive to the acoustic detector, the speaker identity recognizer being arranged for: (1) detecting and distinguishing speech patterns of a plurality of speakers, and (2) enabling the signal to have the first value [only] in response to the speaker having a recognized speech pattern.

16. (amended) The speech recognition system of claim 15 wherein the processing arrangement is arranged for causing the signal to have the first value [only] in response to the speaker having a recognized face matched with a recognized speech pattern of the same speaker.

17. (amended) The speech recognition system of claim 1 wherein the processing arrangement includes a face recognizer connected to be responsive to the visual detector and a speaker identity recognizer connected to be responsive to the acoustic detector, the face recognizer being arranged for detecting and distinguishing the faces of a plurality of speakers, the speaker identity recognizer being arranged for detecting and distinguishing speech patterns of a plurality of speakers, the processing arrangement being arranged for enabling the signal to have the first value [only] in response to the speaker having a recognized face matched with a recognized speech pattern of the same speaker.

18. (amended) A method of recognizing speech utterances of a speaker with an automatic speech recognizer only responsive to acoustic speech utterances of the speaker comprising detecting acoustic energy having a spectrum associated with speech utterances, detecting at least one facial characteristic associated with speech utterances of the speaker, and activating the automatic speech recognizer [only] in response to the detected acoustic energy having a spectrum associated with speech utterances while

Serial No. sn

the at least one facial characteristic associated with speech utterances of the speaker is occurring.

25. (amended) The method of claim 24 further comprising distinguishing the face of the speaker from a plurality of speakers, distinguishing the speech pattern of the speaker from a plurality of speakers, and activating the automatic speech recognizer [only] in response to the speaker having a recognized face matched with a recognized speech pattern of the same speaker.

26. (amended) The method of claim 18 further comprising distinguishing the face of the speaker from a plurality of speakers, distinguishing the speech pattern of the speaker from a plurality of speakers, and activating the automatic speech recognizer [only] in response to the speaker having a recognized face matched with a recognized speech pattern of the same speaker.

27. (amended) The method of claim 26 further including storing: (1) images of the faces of a plurality of speakers, and (2) the speech patterns of the same plurality of speakers during at least one training period; and performing the distinguishing [steps] step by comparing the stored images and speech patterns with images of the face of the speaker and the speech pattern of the speaker.